

A mutational analysis of the two motifs common to adenine methyltransferases

Damion F. Willcock, David T.F. Dryden and Noreen E. Murray

Institute of Cell and Molecular Biology, University of Edinburgh, The Darwin Building, The King's Buildings, Edinburgh EH9 3JR, UK

Communicated by N.E. Murray

All methyltransferases that use S-adenosyl methionine as the methyl group donor contain a sequence similar to (D/E/S)XFXGXG which has been postulated to form part of the cofactor binding site. In N6-adenine DNA methyltransferases there is a second motif, (D/N)PP(Y/F), which has been proposed to play a role similar to the catalytically essential PC motif conserved in all C5-cytosine DNA methyltransferases. We have made a series of amino acid changes in these two motifs in the *EcoKI* N6-adenine DNA methyltransferase. The mutant enzymes have been purified to homogeneity and characterized by physical biochemical methods. The first G is the most conserved residue in motif I. Changing this G to D completely abolished S-adenosyl methionine binding, but left enzyme structure and DNA target recognition unaltered, thus documenting the S-adenosyl methionine binding function of motif I in N6-adenine methyltransferases. Substitution of the N with D, or F with either G or C, in motif II abolished enzyme activity, but left S-adenosyl methionine and DNA binding unaltered. Changes of F to Y or W resulted in partial enzyme activity, implying that an aromatic residue is important for methylation. The substitution of W for F greatly enhanced UV-induced cross-linking between the enzyme and S-adenosyl methionine, suggesting that the aromatic residue is close in space to the methyl-group donor.

Key words: S-adenosyl methionine/DNA methylation/methyltransferase/restriction–modification system

Introduction

DNA modification by the methylation of specific bases occurs in a wide variety of organisms. DNA methylation is both base and sequence specific. In prokaryotes, methylation of either cytosine at the C5 or N4 position, or adenine at the N6 position, is seen as the modification component of restriction–modification (R–M) systems which serves to differentiate self from foreign DNA [see Wilson and Murray (1991), Bickle and Krüger (1993), Heitman (1993), and Roberts and Halford (1993) for recent reviews]. Adenine methylation in prokaryotes has a wide variety of known roles. It can serve to distinguish template DNA from the newly synthesized strand, and it can influence such basic functions as the initiation of DNA replication and the control of gene expression (see

Barras and Marinus, 1989; Noyer-Weidner and Trautner, 1993). In eukaryotes, the C5 position on cytosine residues is the target of methylation and the consequent modification has been implicated in a plethora of important functions (see Bird, 1986; Klimasauskas *et al.*, 1994; Kumar *et al.*, 1994).

The amino acid sequences of many methyltransferases (MTases) have been deduced from their coding sequences and comparative analyses have shown the C5-cytosine MTases to share an ordered set of sequence motifs that alternate with non-conserved regions (Klimasauskas *et al.*, 1989; Pósfai *et al.*, 1989; Kumar *et al.*, 1994). Five or six motifs are well conserved and are presumed to indicate sites for chemistry that is common to all C5-cytosine MTases. Motif I, (D/E/S)XFXGXG, is found in both cytosine and adenine MTases, and is postulated to be the binding site for the cofactor S-adenosyl methionine (AdoMet). In C5-cytosine MTases, motif IV includes an invariant PC dipeptide known to be part of the catalytic site; a covalent linkage between the cysteine and carbon-6 of cytosine has been demonstrated (Chen *et al.*, 1991). The recent co-crystal structure of the *HhaI* MTase plus AdoMet confirms that motif I interacts with the cofactor, and shows this AdoMet binding site to be in close proximity to the catalytic site (Cheng *et al.*, 1993; Klimasauskas *et al.*, 1994).

Much less is known about the adenine MTases. They include one motif that is commonly aligned with motif I and predicted to be the AdoMet binding site. In N6-adenine MTases, this motif is less well conserved than in C5-cytosine MTases and frequently has alanine rather than phenylalanine (Klimasauskas *et al.*, 1989). The second motif characteristic of N6-adenine MTases and N4-cytosine MTases, (N/D/S)PP(Y/F), is well conserved and has been aligned with the PC motif of the C5-cytosine MTases (Klimasauskas *et al.*, 1989; Pósfai *et al.*, 1989).

The modification component of the type I R–M system of *Escherichia coli* K-12 (*EcoKI*) is an adenine MTase composed of two methylation (M) subunits and one DNA sequence specificity (S) subunit. The M subunits include the NPPF (motif II) sequence (Loenen *et al.*, 1987) and a less obvious motif I. In *EcoKI* the latter is DPAXGXA, although DPAXGXG has been identified in two other families of type I R–M systems (Sharp *et al.*, 1992).

We have made mutations within each of these conserved motifs and have characterized the resulting enzymes. The NPPF motif is shown to be critical to the catalysis of the methyl transfer, in agreement with recent data from type II MTases (Sugisaka *et al.*, 1989; Guyot *et al.*, 1993; Kossykh *et al.*, 1993). Although mutations in the NPPF of *EcoKI* do not affect AdoMet binding, this motif is sufficiently close to the binding site to permit cross-linking with the methyl donor AdoMet.

A change of G→D in motif I abolishes cofactor binding,

leaving an enzyme that can bind specifically to its target sequence but, as expected, one in which the binding affinity is no longer enhanced by the presence of the cofactor.

Results

Isolation of mutants

The MTase component of *Eco*KI comprises polypeptides encoded by two genes, *hsdM* and *S* (Host Specificity of DNA). Both motifs characteristic of adenine MTases are in HsdM (the M polypeptide). The mutations in motifs I and II, and the sequences of oligonucleotides used to generate them, are shown in Figure 1.

Modification phenotypes of the *hsdM* mutants

Modification *in vivo* by mutant MTases was assessed from the effectiveness with which they protected phage λ from restriction by the *Eco*KI system. Each λ *hsdMS* phage encoding a mutant MTase was propagated in a host deleted for the *hsd* genes (NM679), and the titre of the resulting lysate was determined on an r_K^+ strain relative to an r_K^- strain (efficiency of plating, e.o.p.). In these tests, λ *hsdM*⁺*S*⁺ phages had an e.o.p. approaching 1 (3×10^{-1}), indicating almost complete protection against restriction, whereas a derivative deleted for *hsdM* had an e.o.p. of only $1-2 \times 10^{-4}$.

Five of the seven mutants (Figure 1) had phenotypes indistinguishable from that associated with the absence of an *hsdM* gene; the phenotypes of the remaining two indicated MTase activity. For one of these (F269W), an e.o.p. of 10^{-2} was consistently observed. This demonstrates substantial, though incomplete, protection of the λ vector from restriction. The remaining mutant (F269Y) had a wild-type phenotype, as assessed from an e.o.p. of 3×10^{-1} . Therefore, MTase activity is retained when the phenylalanine residue is substituted by either a tyrosine or a tryptophan.

MTase structure

The six mutant MTases that differed from wild type by only one amino acid substitution were purified. All six

MOTIF	SEQUENCES	MUTATION AND AMINO ACID SUBSTITUTION
I	-GACCCGGCGGCAGGTACGGCGG- 173 174 175 176 177 178 179 D P A A G T A	
	C CGGCGGCAG TACGGCGG A	GGT→GAT, G 177 D *
II	-CGCCACTAACCCCGCGTTTGGCAGCGCCGAG- 266 267 268 269 N P P F	
	G G CTAACCCCGCG G GGCAGCGCCGAG I I	TTT→GGG, F 269 G TTT→TGG, F 269 W TTT→TGT, F 269 C
	CCGCCGTATGGCAGC	TTT→TAT, F 269 Y
	CGCCACTGACCCGCC	AAC→GAC, N 266 D
	CGCCACTGACCCCGCGTATGGCAGC	(AAC→GAC, N 266 D, TTT→TAT, F 269 Y

Fig. 1. Mutations in motifs I and II. The nucleotide sequences spanning the coding sequences for both motifs are shown. Codon numbers and amino acids identify the motifs in the M polypeptide of *Eco*KI. Beneath each of the two motifs are the oligonucleotide sequences used to direct mutagenesis. Deviations from wild-type sequences are underlined. *, no example of the mutation GGT→GCT was found despite the degeneracy of the oligonucleotide at this position.

were identical to the wild type in subunit stoichiometry (M_2S_1) and molecular weight (170 000), as determined by the relative staining of the M and S subunits by Coomassie Blue on sodium dodecyl sulphate (SDS)-polyacrylamide gels and elution volume on size-exclusion chromatography columns (not shown). Additionally, the guanidinium chloride (GuCl)-induced unfolding of the mutants, as measured by fluorescence spectroscopy, showed a similar GuCl concentration dependence to the wild type (not shown). The amino acid changes therefore do not significantly perturb the protein structure.

The only deviations noted during the purification of the mutants were for the F269C and G177D mutants. The F269C mutant had a tendency to aggregate slightly over a period of several hours in the absence of glycerol and the yield of protein from the purification was poorer. The G177D mutation produced MTase which was recovered only in the pellet of cell debris produced after the initial centrifugation of sonicated cells. This insolubility was found if the transformed cells were grown and transcription induced at 37 or 30°C, but the mutant protein remained

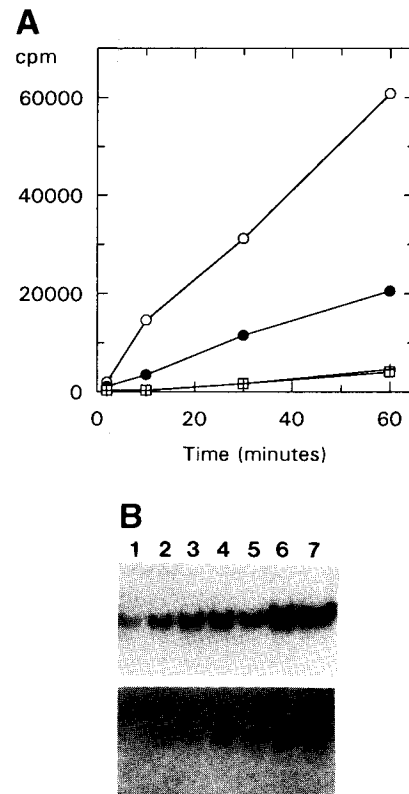


Fig. 2. The methyltransferase activity of the wild-type and mutant enzymes. (A) The rate of methylation of a hemimethylated oligonucleotide determined by scintillation counting of the tritiated DNA. Wild-type MTase (○), F269Y (●), F269W (□) and F269C (+). The concentrations of MTase, AdoMet and DNA were 0.5, 3 and 1.5 μ M respectively. (B) Methylation rate determined by fluorography of the reaction products after electrophoresis on a 10% polyacrylamide gel showing the time course of formation of tritiated DNA. Lanes 1-7 correspond to times of 2, 5, 10, 15, 20, 25 and 30 min for the wild-type (upper panel) and F269W (lower panel). The F269W fluorograph was obtained after 23 days exposure due to the low activity of this enzyme, the wild-type fluorograph after 3 days exposure. The four inactive mutant enzymes gave no fluorographic signal (data not shown). The concentrations of MTase, AdoMet and DNA were 0.1, 3 and 1.5 μ M respectively.

in the supernatant if the cells were grown and induced at the lower temperature of 25°C. The soluble MTase could then be purified in the normal way.

Methyltransferase activity *in vitro*

Only the two mutant enzymes shown to have activity *in vivo* were able to methylate DNA *in vitro*. The *in vitro* methylation rate of hemimethylated DNA, measured by scintillation counting of tritiated DNA bound to an ion-exchange filter, was ~25% of the wild-type rate for the F269Y enzyme, but the rate for the F269W enzyme was indistinguishable from the background level given by an inactive mutant enzyme (Figure 2A). The fluorographic determination of methylation activity proved to be a more sensitive assay, and clearly distinguished between a low methylation activity for the F269W mutant (Figure 2B) and zero activity for the four inactive mutant enzymes. These results reinforce the *in vivo* methylation tests. The fluorographic assay also showed that F269Y and F269W retained the preference of the wild-type enzyme for methylating hemimethylated targets rather than unmodified targets (not shown).

Cofactor binding

The quantitative binding of the AdoMet cofactor was examined by measuring the fluorescence change of 1,8-anilino-naphthalene sulphonic acid (ANS), which binds

weakly to the AdoMet binding site, as AdoMet was added to a solution of the MTase and ANS. The fluorescence of ANS is greater when bound to the MTase than when in solution (Powell *et al.*, 1993). Further confirmation of the AdoMet binding properties was sought by gel filtration of radiolabelled AdoMet–MTase complexes and by UV light-induced cross-linking of radiolabelled AdoMet to the M subunits of the MTase. All of the mutant enzymes except G177D and F269W were identical to the wild type in the above experiments (Figure 3).

The G177D mutant was defective in its affinity for AdoMet, as determined by each of the three methods. The titration of the G177D mutant by AdoMet failed to show any displacement of ANS, even at the highest concentrations of AdoMet (>200 µM). At ~300 µM AdoMet, the ANS fluorescence decreased abruptly and some aggregation of the protein was observed (Figure 3A). The AdoMet binding properties of this enzyme were further checked using gel filtration of a small volume of a radiolabelled AdoMet plus protein mixture to rapidly separate bound AdoMet and free AdoMet. The wild-type and mutant MTases other than G177D showed a peak of radioactivity corresponding to protein-bound AdoMet eluting before a peak of free AdoMet. The G177D mutant AdoMet showed only a peak of free AdoMet, thus confirming that the G177D change essentially abolishes cofactor binding (Figure 3B). The level of UV-induced

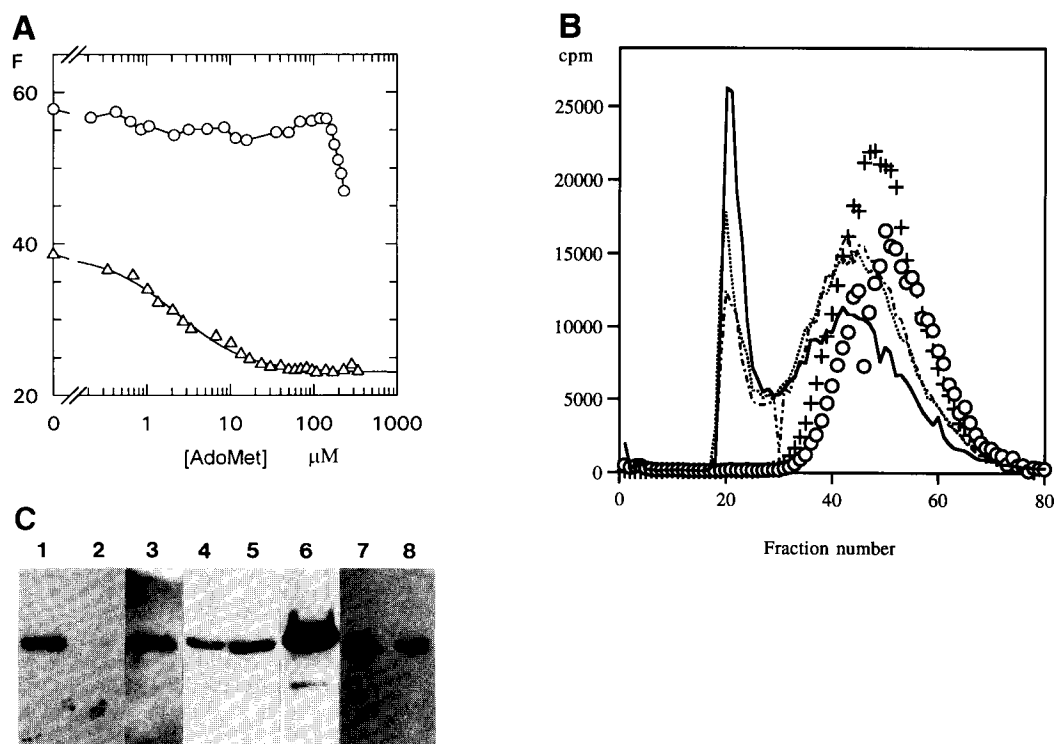


Fig. 3. Cofactor binding by the wild-type and mutant enzymes. (A) The displacement of ANS (50 µM) bound to the MTase (1 µM) and consequent reduction of ANS fluorescence as a function of added AdoMet. The fluorescence emission (arbitrary units) at 480 nm was excited at 395 nm. The wild-type MTase (Δ) shows a 50% change in fluorescence at the K_d of 2.2 µM for AdoMet, but the G177D data (\circ) show no displacement of ANS except at very high AdoMet concentrations. (B) Elution of [3 H]methyl AdoMet with MTase on a Sephadex G25 gel-filtration column. Unbound AdoMet elutes in the later fractions, while protein-bound AdoMet elutes in a sharp peak around fraction 20. AdoMet binds to the wild-type MTase (—), F269C MTase (....) and F269G (— — —) but not to G177D MTase (\circ). The elution of AdoMet in the absence of protein is also shown (+). (C) UV induced cross-linking of [3 H]methyl AdoMet (3 µM) to the M subunits of the MTase (1.4 µM) shown by fluorography of the MTase–AdoMet complexes on a 10% polyacrylamide–SDS gel. Lanes 1 and 8, wild-type; lanes 2–7 are G177D, F269C, N266D, F269Y, F269W and F269G respectively.

cross-linking was very low, consistent with a very weak affinity of the mutant enzyme for the cofactor (Figure 3C, lane 2).

The F269W mutant showed wild-type binding affinity for AdoMet, as determined by the ANS displacement measurements, but also showed an enormously enhanced (~40-fold) UV light-induced cross-linking of the M subunit to the cofactor (Figure 3C, lane 6).

DNA binding

DNA binding was measured for each of the mutant enzymes using the gel-retardation technique in which a fixed concentration of a 32 P end-labelled oligonucleotide is titrated with increasing amounts of protein. Protein-DNA complexes were separated from free DNA on a non-denaturing polyacrylamide gel and the amount of free DNA quantified by autoradiography to give an estimate of binding to an unmodified specific oligonucleotide and to a non-specific oligonucleotide in the presence and absence of AdoMet. The wild-type MTase binds to the specific target with a K_d of 9.0 ± 2.2 and 2.1 ± 0.7 nM in the absence or presence of AdoMet, respectively, and to the non-specific DNA with a K_d of 89.3 ± 13.5 and 43.0 ± 6.6 nM under the same conditions (Powell *et al.*, 1993). All of the mutant MTases, except the G177D mutant, showed DNA affinities similar to those for the wild type (data not shown).

Since the G177D mutant does not bind AdoMet, one predicts that the DNA binding affinity would not be

influenced by the cofactor. This was found to be the case (Figure 4). The mutant enzyme bound the specific and non-specific DNA in the presence or absence of AdoMet with similar affinity to the wild-type enzyme in the absence of AdoMet. The tendency of the MTase to form a second slower migrating non-specific complex at high protein concentration was accentuated in the G177D mutant.

Discussion

We have purified mutant *Eco*KI MTases that have single amino acid changes in either one of the two motifs conserved in all adenine MTases (see Figure 1). Both motifs I and II occur in each of the two M subunits of the trimeric MTase and one molecule of AdoMet binds to each of these subunits (Dryden *et al.*, 1993; Powell *et al.*, 1993). Motif I occurs in the N-terminal domain and motif II forms part of a proteolytically sensitive surface loop linking the N- and C-terminal domains of M (Figure 5). The binding of AdoMet protects the loop from proteolysis, possibly by making the loop close down in a rigid conformation on the protein surface (Cooper and Dryden, 1994).

Our characterization of the mutant enzymes divides them into three groups, comprising one inactive enzyme which does not bind cofactor (G177D in motif I), three inactive enzymes which do bind cofactor (N266D, F269C and F269G in motif II) and two partially active enzymes (F269Y and F269W in motif II). All had the normal,

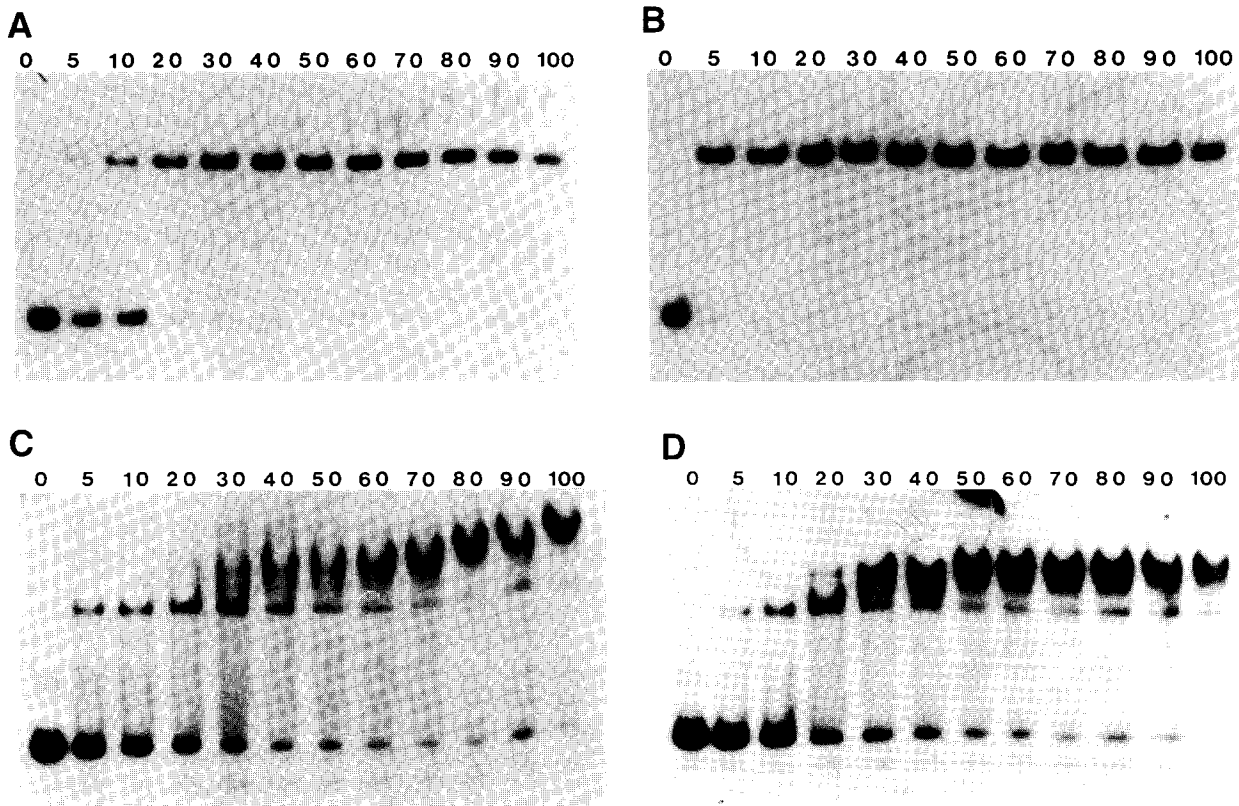


Fig. 4. Retardation of 0.1 nM 32 P end-labelled unmodified oligonucleotide on 5% polyacrylamide gels by the MTase. The protein concentration is shown in nM above each lane. (A) and (B) show retardation by the wild-type MTase in the absence and presence of 100 μ M AdoMet respectively. (C) and (D) show the effect of G177D MTase in the absence and presence of AdoMet, respectively. A second retarded non-specific complex becomes visible at the highest protein concentrations.

stable M_2S_1 trimeric structures and could bind specifically to the *EcoKI* DNA target. This implies that none of the amino acid changes causes a gross change in the protein tertiary structure.

Our experiments indicate that the G177D mutation in *EcoKI* abolishes enzyme activity by dramatically reducing the affinity of the M subunits for the AdoMet cofactor. The loss of AdoMet binding, and probably the sensitivity of this mutant to the cell growth temperature, can therefore be attributed to a localized conformational disturbance caused by the replacement of the small glycine residue with a larger, charged aspartic acid residue which is much more restricted in its conformation. This result indicates that motif I in *EcoKI* and, by sequence comparisons, in other N6-adenine MTases, forms a crucial part of the cofactor binding site. We have isolated two other mutations in motif I, D173G and P174S, which display r^-m^- and r^-m^{ls} phenotypes, respectively (V.A.Barcus and A.Daniel, unpublished results, this laboratory). The D173G change results in an insoluble enzyme at either 25 or 37°C. The phenotypes of these two mutants are consistent with a defect in AdoMet binding.

Glycine 177 in the M subunit is part of the sequence DPAAGTA within the consensus sequence of motif I (D/E/S)XAXGXG, found in N6-adenine and N4-cytosine

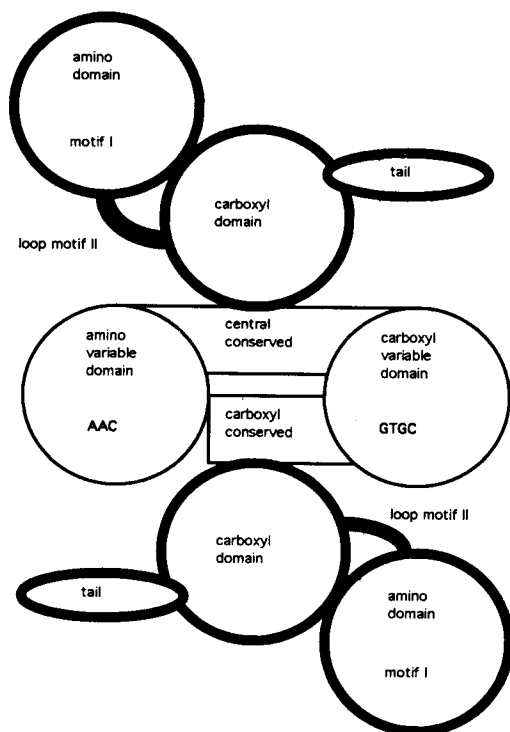


Fig. 5. A sketch of the domain structure of *EcoKI* MTase as derived from limited proteolysis experiments (Cooper and Dryden, 1994). Each M subunit (529 amino acids, bold outline) has an N-terminal domain (~260 amino acids) containing motif I, a short loop (~15 amino acids) containing motif II, a C-terminal domain (~150 amino acids) contacting the S subunit and a tail (~100 amino acids) which is sensitive to the methylation of the DNA target. The S subunit (463 amino acids, light outline) contains two target recognition domains and recognizes the sequence AAC(N)₆GTGC: the N-terminal domain (~150 amino acids) recognizes the AAC sequence, the C-terminal domain (~180 amino acids) recognizes GTGC, these domains being separated by small amino acid regions (~30 and ~80 amino acids) conserved between S subunits of different type I systems.

MTases (Klimasauskas *et al.*, 1989; Sharp *et al.*, 1992). The underlined glycine is the most conserved residue in the motif. Sequences similar to this motif can be found in all MTases which utilize AdoMet and it should be possible to define the structure of the motif in *EcoKI* by comparison with known structures.

The recent elucidation of the tertiary structure of the *HhaI* C5-cytosine MTase bound to AdoMet or to S-adenosyl homocysteine plus DNA confirms the role of motif I in AdoMet binding in C5-cytosine MTases (Cheng *et al.*, 1993; Klimasauskas *et al.*, 1994), postulated by Wilke *et al.* (1988) from a mutational analysis of the SPR MTase. The first glycine residue in the motif in *HhaI* forms a sharp turn between a β strand and an α helix. The overall structure of the AdoMet binding domain containing motif I is similar to the nucleotide binding domain (the so-called Rossman fold) (Branden and Tooze, 1991; Creighton, 1993). This domain structure and motif I have also been found in catechol *O*-methyltransferase which requires AdoMet as cofactor (Vidgren *et al.*, 1994).

Our secondary structure predictions (D.T.F.Dryden, unpublished results) for the M subunits of type I R-M systems suggest that motif I and the surrounding sequence form the same β strand-turn- α helix structure in *EcoKI* N6-adenine MTase as in the two known MTase structures. Our G177D substitution would therefore hinder the formation of the turn in the secondary structure and the formation of the correct contacts with AdoMet. In addition, given that two different MTases share an AdoMet binding domain structure similar to a common nucleotide binding domain, it is probable that *EcoKI* also contains similar domains. Secondary structure prediction indicates that the β strand- α helix- β strand motifs which make up the Rossman folding unit occur in the domains of the M subunit (D.T.F.Dryden, unpublished results).

Our other amino acid changes are in motif II, (N/D/S)PP(F/Y), which is found in N6-adenine and N4-cytosine MTases. This motif, first reported for Dam MTases (Hattman *et al.*, 1985), can be aligned with the catalytically essential PC motif found in all C5-cytosine MTases when the neighbouring amino acids are also considered (Klimasauskas *et al.*, 1989). In the *HhaI* structure, the PC motif is in a long surface loop of the protein whose conformation changes dramatically upon DNA binding (Cheng *et al.*, 1993; Klimasauskas *et al.*, 1994). The alignment of motif II with the PC motif of C5-cytosine MTases is supported by the following experiments which indicate a surface position for the motif.

The limited proteolysis of the *EcoKI* MTase shows that motif II is highly sensitive to proteases except when AdoMet is bound (Cooper and Dryden, 1994). This implies that AdoMet binding alters the conformation of motif II so that it is not so exposed to the solvent. A surface position is also suggested by secondary structure predictions (D.T.F.Dryden, unpublished results). The Mnt repressor protein contains the sequence DHPF, but when this is changed to DPPF the target sequence for the mutant repressor includes N6 methylated adenine, implying a direct interaction of the amino acid motif and the N6 methylated adenine (Vershon *et al.*, 1985; Chandrasegaran and Smith, 1988), and lastly a preliminary report on the structure of *TaqI* N6-adenine MTase indicates a surface loop position for motif II (Anderson, 1993).

Three of our mutations (F269C, F269G, N266D) in motif II abolished MTase activity, but left substrate binding and specificity unaltered. The N266D change, along with similar changes in the Dam MTase (D→N, S or G) (Guyot *et al.*, 1993; Kosykh *et al.*, 1993) and *FokI* MTase (D→N, G or A) (Sugisaka *et al.*, 1989), shows that this first residue in motif II cannot be substituted, even with other amino acids within the consensus sequence. In our double mutant, the F269Y change, which by itself does not abolish activity, fails to compensate for the N266D change and the enzyme is inactive.

Only two of our motif II amino acid changes produce active enzyme. The F269Y and F269W mutants showed normal substrate binding and the same strong preference for methylating hemimethylated targets rather than unmethylated targets as the wild-type MTase (Suri and Bickle, 1985; Dryden *et al.*, 1993), albeit at a slower overall methylation rate. F269Y possessed ~25% and F269W <5% of wild-type activity. The low activity of the F269W mutation may be due to some steric hindrance by the large indole ring in the reaction. The extremely strong cross-linking of AdoMet to the F269W mutant suggests that the tryptophan is the site of cross-linking. If this assumption is correct, then the tryptophan—and by extrapolation the normal phenylalanine residue—are very close to AdoMet in the MTase–AdoMet complex. Therefore, although motif II appears to have little influence on the binding affinity for AdoMet in *EcoKI*, it is clearly in the vicinity of the binding site. It is also possible, but we believe more unlikely given tryptophan's photoreactivity, that the indole ring pushes the side chain of another amino acid into contact with the AdoMet. Changing the proline residues in motif II in the T4 Dam MTase (Kosykh *et al.*, 1993) or *E. coli* Dam MTase (Guyot *et al.*, 1993) reduced or abolished AdoMet binding, respectively, indicating that the motif, as well as being essential for activity, can play a part in substrate binding.

Recently, the gene sequence of *VspI*, an N6-adenine MTase of a *Vibrio* species, was determined (Degtyarev *et al.*, 1993). The amino acid sequence of motif II was found to be NPPW, thus supporting our results for the F269W mutant and indicating that the consensus sequence of motif II can be extended to (N/D/S)PP(F/Y/W).

In conclusion, our results on *EcoKI* N6-adenine MTase provide mutational evidence that motif I is part of the AdoMet binding site in the M subunit, in agreement with the recent crystallographic structures of C5-cytosine *HhaI* MTase and catechol *O*-methyltransferase (Cheng *et al.*, 1993; Klimasauskas *et al.*, 1994; Vidgren *et al.*, 1994).

Changes in motif II did not affect substrate binding and specificity, but some abolished activity, suggesting a direct role in catalysis. The asparagine residue cannot be replaced by the aspartic acid residue present in other adenine MTases, thus separating these MTases into two groups. Changes of the phenylalanine produced enzymes which were active only if the substitution was an aromatic residue, therefore implying that the aromatic residue is essential for methyl group transfer.

Materials and methods

Bacteria, phage and plasmids

Three contiguous chromosomal genes, *hsdR*, *M* and *S*, encode *EcoKI*. A λ phage (λ NM1048), including *hsdS*, *hsdM* and the major part of

hsdR, has been described previously (Sain and Murray, 1980). In the present experiment we also use λ NM1065 (Sain and Murray, 1980), a deletion derivative of λ NM1048, made by the excision of a *Bam*HI fragment that includes the 3' end of *hsdR* and most of *hsdM* (see Kelleher *et al.*, 1991). All three *hsd* genes are needed to encode the restriction endonuclease and confer the $r_K^+ m_K^+$ phenotype, but *hsdM* and *S* are sufficient for modification ($r_K^- m_K^-$). The *E. coli* strains used in this paper, other than the *mutL* derivative of BMH 71–81 (Kramer *et al.*, 1984), are deleted for at least that part of *hsdM* that is the target for mutagenesis.

NM522 (Gough and Murray, 1983), the host for M13, lacks *hsdM* and *hsdS*. Alternative *hsd* Δ hosts for plasmids and λ phage were NM654 (Kelleher *et al.*, 1991), a derivative of C600 with the same deletion as λ NM1065, and NM679, a derivative of W3110 in which a deletion removes the entire *hsd* region. C600 was used as a standard $r_K^+ m_K^+$ strain (Appleyard, 1954).

The substrate for site-directed mutagenesis was a 1.4 kb *Sma*I–*Bam*HI fragment from λ NM1048 including the 5' end of *hsdM* cloned in M13mp19 (Loenen *et al.*, 1987). Mutated *hsdM* sequences were excised on a 1.4 kb *Bam*HI fragment by making use of the flanking *Bam*HI target in the polylinker of M13mp19 (Yanisch-Perron *et al.*, 1985). The coding sequence for the MTase, but not the HsdR polypeptide, was regenerated by ligating this *Bam*HI fragment to λ NM1065 DNA cut with *Bam*HI. The resulting phages were recovered on either NM654 or NM679. No functional HsdR polypeptide was present in either host, so that a mutation that inactivated the MTase, but not the endonuclease activity, would not be lethal.

A derivative of pBR322, including the *Bam*HI fragment deleted from λ NM1065, was used to probe for λ derivatives that had acquired the 1.4 kb insert.

The expression vector pJF118HE (Furst *et al.*, 1986) was used to overexpress both wild-type and mutant MTases. The wild-type derivative (pJFMS) was available (Dryden *et al.*, 1993); the others were made in the same way as pJFMS by transferring a 7.8 kb *Sma*I–*Eco*RI fragment excised from the appropriate derivatives of λ NM1065.

Media and microbiological methods

Media and general methods have been described previously (Murray *et al.*, 1977). The modification phenotypes of *hsdM* phages were assessed after multiple rounds of infection of the *hsd* Δ strain NM679. The e.o.p. of these phages on C600 ($r_K^+ m_K^+$) relative to NM679 ($r_K^- m_K^-$) reflects the level of *in vivo* methylation.

Preparation, manipulation and recovery of DNA

The methods used were those described by Midgley and Murray (1985).

Site-directed mutagenesis

DNA of the M13mp19 derivative, including the relevant *hsdM* sequence (see Figure 1), was used as template for mutagenesis directed by synthetic oligonucleotides (Zoller and Smith, 1983). Oligonucleotides were synthesized by OSWEL DNA (University of Edinburgh). The template, oligonucleotide sequences and changes produced are shown in Figure 1. Infective genomes were recovered in the *mutL* strain BMH 71–81, and plaques from phage containing mutations were identified initially by differential hybridization and then by nucleotide sequence. The 1.4 kb *Bam*HI fragment was excised from Rf DNA prepared in NM522 and transferred to λ NM1065. λ phages with inserts were identified by hybridization, DNA was made and the orientation of the *Bam*HI fragment determined from the analysis of restriction digests. Modification phenotypes of phages with the fragment in the correct orientation were determined and the *Bam*HI fragment was returned to M13mp19 and its complete nucleotide sequence checked to confirm that the site-directed mutation was the only difference from the wild-type sequence.

DNA sequencing

The nucleotide sequence of the 1.4 kb *Bam*HI fragment from each mutant was determined by the modification of the dideoxy chain termination method using Sequenase (Tabor and Richardson, 1987). The nucleotide sequences were compared with that of the wild-type *hsdM* gene (Loenen *et al.*, 1987) and nucleotide substitutions identified.

Protein purification and characterization

NM679, freshly transformed by the appropriate plasmid, was used as the source of enzyme. The MTase was purified as described by Dryden *et al.* (1993) and all mutant enzyme preparations were homogeneous, as judged by Coomassie Blue staining of SDS–10% polyacrylamide gels (Laemmli, 1970). Owing to the insolubility of the G177D MTase in

cells grown at 37 or 30°C, these cells were grown at 25°C overnight and induced for 8 h at 25°C. Protein concentrations were measured by UV absorption using an extinction coefficient of 0.842 cm⁻¹ at 280 nm for a 1 mg/ml solution for all mutants except F269Y and F269W, which had extinction coefficients of 0.861 and 0.912 cm⁻¹.

Protein molecular weight, stability, AdoMet binding and DNA binding experiments were performed as described previously (Dryden *et al.*, 1993; Powell *et al.*, 1993).

AdoMet binding was also measured by gel filtration of [³H]methyl AdoMet (Amersham) plus MTase samples. Fifty microlitres of 12 µM MTase plus 1 µl of 12 µM [³H]methyl AdoMet samples were applied to a 5 mm diameter × 14.6 cm long column of Sephadex G25 media (Pharmacia) equilibrated in 20 mM Tris, 20 mM MES, 200 mM NaCl, 10 mM MgCl₂, 7 mM β mercaptoethanol, 0.1 mM EDTA (pH 8). Fractions (~100 µl) of two drops each were collected, 2.5 ml of Ecoscint (National Diagnostics) were added to each and the ³H counted in a Beckmann LS7000 scintillation counter.

The rate of methylation of DNA was measured by following the transfer of [³H]methyl groups to target DNA and subsequent scintillation counting of the tritiated DNA retained by DE81 ion-exchange paper (Whatman) (Dryden *et al.*, 1993) or by the following method. A 40 µl reaction containing 1.5 µM DNA duplex, 20–150 nM MTase and 3 µM [³H]methyl AdoMet in 100 mM Tris, 100 mM NaCl, 7 mM MgCl₂ (pH 7.9) was made. Amounts of 5 µl were withdrawn at appropriate times and the reaction stopped by heating the sample to 65°C. Reaction products were separated on a 10% polyacrylamide non-denaturing gel which was subsequently fixed in 10% acetic acid, 10% methanol for 30 min. The gel was then soaked in Amplify (Amersham) for 30 min, dried and subjected to fluorography to show the time course of tritiation of the DNA duplex. The DNA substrate for methylation activity and DNA binding was of the sequence 5'-TGTCTAGATATCGGCCTAA*CC-ACGTGGTGCCTACGAGCTCAGGCG-3' and its complementary strand. The recognition sequence of EcoKI is underlined (Kan *et al.*, 1979). One of the duplexes contained a methylated base at the * position to give a hemimethylated substrate.

Acknowledgements

We thank Mr L.P.Cooper, Ms K.E.Munn and Dr L.M.Powell for their assistance with the assays, mutagenesis and gel retardation experiments, Mrs Isobel Black for her help in the preparation of the manuscript, and Graham Brown and Frank Johnston for photography. This work was supported by the Medical Research Council and the Science and Engineering Research Council who awarded a studentship to D.F.W. This is a publication from the Edinburgh Centre for Molecular Recognition.

References

- Anderson, J.E. (1993) *Curr. Opin. Struct. Biol.*, **3**, 24–30.
 Appleyard, R.K. (1954) *Genetics*, **39**, 440–459.
 Barras, F. and Marinus, M.G. (1989) *Trends Genet.*, **5**, 139–143.
 Bickle, T.A. and Krüger, D.H. (1993) *Microbiol. Rev.*, **57**, 434–450.
 Bird, A.P. (1986) *Nature*, **321**, 209–213.
 Branden, C. and Tooze, J. (1991) *Introduction to Protein Structure*. Garland Publishing, New York, pp. 18–19.
 Chandrasegaran, S. and Smith, H.O. (1988) In Sarma, R.H. and Sarma, M.H. (eds), *Structure and Expression from Proteins to Ribosomes*. Adenine Press, New York, pp. 149–156.
 Chen, L., MacMillan, A.M., Chang, W., Ezaz-Nikpay, K., Lane, W.S. and Verdine, G.L. (1991) *Biochemistry*, **30**, 11018–11025.
 Cheng, X., Kumar, S., Pósfai, J., Pflugrath, J.W. and Roberts, R.J. (1993) *Cell*, **74**, 299–307.
 Cooper, L.P. and Dryden, D.T.F. (1994) *J. Mol. Biol.*, **236**, 1011–1021.
 Creighton, T.E. (1993) *Proteins: Structures and Molecular Properties*. 2nd edn. W.H. Freeman & Co., New York, pp. 250–251.
 Degtyarev, S.Kh., Prikhod'ko, E.A., Prikhod'ko, G.G. and Krasnykh, V.N. (1993) *Nucleic Acids Res.*, **21**, 2015.
 Dryden, D.T.F., Cooper, L.P. and Murray, N.E. (1993) *J. Biol. Chem.*, **268**, 13228–13236.
 Furst, J.P., Pansegrau, W., Frank, R., Blocker, H., Scholz, P., Bagdasarian, M. and Lanka, E. (1986) *Gene*, **48**, 119–131.
 Gough, J.A. and Murray, N.E. (1983) *J. Mol. Biol.*, **166**, 1–19.
 Guyot, J.-B., Grassi, J., Hahn, U. and Guschlbauer, W. (1993) *Nucleic Acids Res.*, **21**, 3183–3190.
 Hattman, S., Wilkinson, J., Swinton, D., Schlagman, S., MacDonald, P.M. and Mosig, G. (1985) *J. Bacteriol.*, **164**, 932–937.

- Heitman, J. (1993) In Setlow, J.K. (ed.), *Genetic Engineering*. Plenum Press, New York, pp. 57–108.
 Kan, N.C., Lautenberger, J.A., Edgell, M.H. and Hutchison, C.A., III (1979) *J. Mol. Biol.*, **130**, 191–209.
 Kelleher, T.E., Daniel, A.S. and Murray, N.E. (1991) *J. Mol. Biol.*, **221**, 431–440.
 Klimasauskas, S., Timinskas, A., Menkevicius, S., Butkiene, D., Butkus, V. and Janulaitis, A. (1989) *Nucleic Acids Res.*, **17**, 9823–9831.
 Klimasauskas, S., Kumar, S., Roberts, R.J. and Cheng, X. (1994) *Cell*, **76**, 357–369.
 Kossykh, V.G., Schlagman, S.L. and Hattman, S. (1993) *Nucleic Acids Res.*, **21**, 3563–3566.
 Kramer, B., Kramer, W. and Fritz, H.-J. (1984) *Cell*, **38**, 879–887.
 Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Pósfai, J., Roberts, S.J. and Wilson, G.G. (1994) *Nucleic Acids Res.*, **22**, 1–10.
 Laemmli, U.K. (1970) *Nature*, **227**, 680–685.
 Loenen, W.A.M., Daniel, A.S., Braymer, H.D. and Murray, N.E. (1987) *J. Mol. Biol.*, **198**, 159–170.
 Midgley, C.A. and Murray, N.E. (1985) *EMBO J.*, **4**, 2695–2703.
 Murray, N.E., Brammar, W.J. and Murray, K. (1977) *Mol. Gen. Genet.*, **150**, 53–61.
 Noyer-Weidner, M. and Trautner, T.A. (1993) In Jost, J.P. and Sahiz, H.P. (eds), *DNA Methylation: Molecular Biology and Biological Significance*. Birkhäuser Verlag, Basel, pp. 39–108.
 Pósfai, J., Bhagwat, A.S., Pósfai, G. and Roberts, R.J. (1989) *Nucleic Acids Res.*, **17**, 2421–2435.
 Powell, L.M., Dryden, D.T.F., Willcock, D.F., Pain, R.H. and Murray, N.E. (1993) *J. Mol. Biol.*, **234**, 60–71.
 Roberts, R.J. and Halford, S.E. (1993) In Linn, S.M., Lloyd, R.S. and Roberts, R.J. (eds), *Nucleases*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 35–88.
 Sain, B. and Murray, N.E. (1980) *Mol. Gen. Genet.*, **180**, 35–46.
 Sharp, P.M., Kelleher, J.E., Daniel, A.S., Cowan, G.M. and Murray, N.E. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 9836–9840.
 Sugisaka, H., Kita, K. and Takanami, M. (1989) *J. Biol. Chem.*, **264**, 5757–5761.
 Suri, B. and Bickle, T.A. (1985) *J. Mol. Biol.*, **186**, 77–85.
 Tabor, S. and Richardson, C.C. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 4767–4771.
 Vershon, A.K., Youderian, P., Weiss, M.A., Susskind, M.M. and Sauer, R.T. (1985) In Calendar, R. and Gold, L. (eds), *Sequence Specificity in Transcription and Translation*. Alan R. Liss, New York, pp. 209–218.
 Vidgren, J., Svensson, L.A. and Liljos, A. (1994) *Nature*, **368**, 354–358.
 Wilke, K., Rauhut, E., Noyer-Weidner, M., Lauster, R., Powlek, B., Behrens, B. and Trautner, T.A. (1988) *EMBO J.*, **7**, 2601–2609.
 Wilson, G.G. and Murray, N.E. (1991) *Annu. Rev. Genet.*, **25**, 585–627.
 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, **33**, 103–119.
 Zoller, M.J. and Smith, M. (1983) *Methods Enzymol.*, **100**, 468–500.

Received on May 13, 1994; revised on June 6, 1994